



UNIVERSIDADE DA CORUÑA

Grado en Biología.

Memoria del Trabajo de Fin de Grado.

Revisión bibliográfica / Investigación: O lastre xenético procedente do Neanderthal nas poboacións do home moderno.

Revisión bibliográfica / Investigación: El lastre genético procedente del Neanderthal en las poblaciones del hombre moderno.

Literature review / Research: The genetic load from the Neanderthal in modern human populations.

Lucía Álvarez Rodríguez.

Septiembre 2018.

Director académico:

ÍNDICE.

1. Resumen / Summary.
2. Palabras clave/ Key words.
3. Introducción.
4. Objetivos.
5. Material y métodos.
6. Resultados.
7. Discusión / Discussion.
8. Conclusiones / Conclusions.
9. Bibliografía.

1. Resumen:

El ser humano no ha estado solo a lo largo de la historia. Otras especies de homínidos (alguna todavía por descubrir) se han cruzado con la nuestra. Y esta huella genética ha sido crucial en la supervivencia o extinción de muchas de ellas, destacando la especie humana, ya que es la única que persiste.

Hoy en día la mayoría de esas huellas de introgresión son perjudiciales, y por lo tanto deben considerarse un lastre genético para las poblaciones contemporáneas. En este trabajo haremos un estudio completo de 9 alelos derivados del cruce con el Neanderthal, además de plantearnos por qué todavía no han sido eliminados.

Summary:

The human being has not been alone throughout the history. Other species of hominids (one still for discovering) have been crossed with ours. And this genetic mark has been crucial in the survival or extinction of many of them, highlighting the human species, since it is the only one that persists.

Nowadays most of those introgression marks are harmful, and must therefore be considered a genetic ballast for the contemporary populations. In this work we will do a complete study of 9 alleles derived from the road junction with the Neanderthal, besides considering why they have not been eliminated yet.

2. Palabras clave:

SNPs, enfermedades, genética poblacional, selección, deriva genética, alelos derivados, introgresión.

Key words: SNPs, diseases, poblational genetic, selection, genetic drift, derived alleles, introgression.

3. Introducción.

Nuestra especie apareció hace 200 mil años, y aunque sólo nosotros hemos llegado a sobrevivir, existían otros homínidos.

Hasta hace relativamente poco tiempo compartíamos Europa y oeste de Asia con los Neanderthales, Asia con los Denisovanos, y con los 'hobbits' en la isla indonesia de Flores.

De los que más conocemos son los Neanderthales, los cuales estuvieron más de 200 mil años en Europa, ya muy bien adaptados. Se extinguieron por el rápido cambio climático,

debido a que la caza del ser humano moderno era mucho más eficiente que la suya, contando además con un juego genético menos afortunado. Pero convivieron con nosotros hasta hace unos 40 mil años, y se produjeron cruces entre especies (Hogenboom, 2015).

Pero en el caso de los Neanderthales, comparten más variantes con los no africanos que con africanos. Esto se debe al cruce que tuvieron con los seres humanos modernos cuando salieron fuera de África.

Una posibilidad que se baraja ahora mismo es que el último contacto con un flujo de genes de origen Neanderthal haya sido no con esta misma especie, sino con una población arcaica relacionada con éstos.

Se sospecha por la relación periodo estimado del cruce - yacimientos arqueológicos encontrados, que el lugar de la mezcla haya podido darse en Eurasia occidental (Sankararaman, Patterson, Li, Pääbo, & Reich, 2012).

Recientemente, gracias a análisis de datos de secuencia de ADN se ha sabido más acerca de los orígenes moleculares del ser humano.

Se cree que el flujo de genes desde el Neanderthal no ha sido nunca una gran contribución, pero ha sido compatible con el recambio de genes a gran escala desde la evolución del hombre arcaico al hombre moderno, de ahí su permanencia en el tiempo. El humano moderno no ha sido la única especie que ha tenido bajos niveles de mezcla con otras especies, sino que se ha demostrado en todos los primeros grupos humanos un ligero flujo de genes, aunque no hayan sobrevivido (Kelso & Prüfer, 2014).

De hecho, numerosos estudios han revelado que la hibridación entre especies durante la evolución humana han sido recurrentes, aunque en el caso del ser humano moderno, la importancia genotipo - fenotipo continúa debatiéndose, aunque es innegable que ha sido una gran herramienta de adaptación y variabilidad de la especie (Ackermann, Mackay, & Arnold, 2016).

Se sabe, además, de la existencia de una red de poblaciones, unidas por un flujo genético, que aunque imitado, ha sido también intermitente, que implica múltiples linajes de homínidos, todos extintos menos el ser humano moderno, el cual contó con una introgresión de alelos que le facilitaron la adaptación a cualquier desafío ambiental. Se propagó por todo el planeta y de encontró con numerosas especies de homínidos, aunque nunca llegó a existir un `híbrido` como tal, sí hubo un reemplazo de loci. Podemos decir que somos la única subespecie que ha reemplazado a todas las demás, incluido cualquier tipo de híbrido.

Y es que el éxito del ser humano moderno puede deberse a unas dotes cognitivas mucho

más complejas, fruto de nuestro genotipo original.

Un dato relevante que ha surgido a partir del análisis de ADN mitocondrial, es que sólo se encontró progenie de ser humano moderno en casos donde la madre también lo fuese.

Es decir, los hombres eran los portadores de genes foráneos, y sólo en este caso sobrevivían, transmitiendo alelos relacionados con adaptación ecológica, pero nunca cognitiva, por ello se entendían bien en grupos de ser humano moderno (Varki, 2016).

4. Objetivos.

1. Describir el estado actual del conocimiento sobre el lastre genético procedente del Neanderthal en las poblaciones del hombre moderno.
2. Identificar alelos en el genoma humano que proceden de introgresión del Neanderthal.
3. Localizar evidencias de asociación de éstos con patologías humanas.
4. Detectar huellas de selección sobre estas regiones del genoma del hombre moderno.

5. Material y métodos.

Este estudio es tanto bibliográfico como de investigación, ya que no se ha limitado a reclutar bibliografía sino también a escrutar resultados propios en buscadores genómicos. Como hemos dicho, escogimos diferentes SNPs (ilustración 1), concretamente los correspondientes al artículo (Patterson et al., 2014).

rs id	Coordinates	Derived allele	Derived allele frequency (%)		Phenotype
			Europeans	East Asians	
rs12531711	7:128,617,466	G	10.03	0.17	Systemic lupus erythematosus, Primary biliary cirrhosis Smoking behavior Crohn's disease Optic disc size Interleukin-18 levels Crohn's disease
rs3025343	9:136,478,355	A	8.44	0.00	
rs7076156	10:64,415,184	A	26.52	8.74	
rs12571093	10:70,019,371	A	16.35	14.86	
rs1834481	11:112,023,827	G	21.50	0.35	
rs11175593	12:40,601,940	T	1.98	3.32	Type-2 Diabetes
rs75493593	17:6,945,087	T	1.85	12.06	
rs75418188	17:6,945,483	T	1.85	11.54	
rs117767867	17:6,946,330	T	1.85	11.54	

Ilustración 1: detalle de la tabla de SNPs.

Aquí tenemos un ejemplo de 9 SNPs, cada uno representa un lugar donde una base ha sido derivada del cruce con el Neanderthal. En nuestro estudio confirmaremos cual era la base ancestral y la derivada, sus frecuencias, las consecuencias fenotípicas, su localización respecto al gen y genes próximos, además de su posible actividad.

El procedimiento ha sido el mismo para todos los SNPs, por lo que, para evitar extender demasiado este apartado, mostraré el método del primero en la lista, entendiendo así que se ha repetido en todos los demás para obtener los datos pertinentes.

Primeramente visitamos la web <https://www.ensembl.org/index.html> e introducimos el id de nuestro rs en el buscador, donde también podremos seleccionar el ser humano como especie.

· SNP rs12531711.

Ahora nos saldrá una ventana de resultados (ilustración 2), comprobamos que corresponda con el rs que hemos introducido, y pinchamos sobre él.

The screenshot shows the Ensembl genome browser interface. At the top, there's a navigation bar with links like BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. A search bar on the right contains the text 'Search all species...'. Below the navigation bar, there's a 'New Search' section. On the left, there are filters for 'Restrict category to:' (set to Variant), 'Restrict species to:' (set to Human), 'Per page:' (set to 10), 'Layout:' (set to Standard), and a 'Tip:' section. The main search area shows the input 'rs12531711' and a search button. Below the search bar, it says '1 results match rs12531711'. A link 'Did you mean...' is visible. The search result for 'rs12531711 (Human Variant)' is displayed, showing a list of associated phenotypes and diseases, including 'Lupus Erythematosus', 'Systemic lupus erythematosus', 'Primary biliary cirrhosis', 'Phenotype ontologies', 'Leukopenia', 'Alopecia', 'Splenomegaly', 'Xerostomia', 'Hematuria', 'Hepatomegaly', 'thrombocytopenia', 'Pulmonary hypertension', 'Conjunctival telangiectasia', 'Coronary artery disease', 'Autoimmunity', 'Skin rash', 'Cutis marmorata', 'Myositis', 'Antinuclear antibody positivity', 'Abnormality of the endocardium', 'Increased intracranial pressure', 'Eczema', 'Verrucae', 'biliary liver cirrhosis', 'Cranial nerve paralysis', 'Telangiectasia of the skin', 'Abnormal tendon morphology', 'Arterial thrombosis', 'Subcutaneous hemorrhage', 'Renal insufficiency', 'Memory impairment', 'Abnormality of the heart valves', 'Hallucinations', 'Abnormal blistering of the skin', 'systemic lupus erythematosus', 'Myalgia', 'Inflammation of the large intestine', 'Hemolytic anemia', 'Hypermelanotic macule', 'Lymphopenia', 'Antiphospholipid antibody positivity', 'Increased antibody level in blood', 'Abnormal pyramidal signs', 'Vasculitis', 'Pleuritis', 'Pulmonary infiltrates', 'Aseptic necrosis', 'diarrhea', 'Nephritis', 'Recurrent respiratory infections', 'Biliary cirrhosis', 'Restrictive ventilatory defect', 'Peripheral neuropathy', 'Thrombophlebitis', 'Sleep disturbance', 'Chest pain', 'Cerebral palsy', 'Cutaneous photosensitivity', 'Fatigable weakness', 'rheumatoid arthritis', 'Edema of the lower limbs', 'Cerebral ischemia', 'Seizures', 'Pancreatitis', 'Lymphadenopathy', 'Cellulitis', 'arthralgia', 'Bone marrow hypocellularity', 'Retinopathy', 'Dry skin', 'Autosomal dominant inheritance', 'Urticaria'. At the bottom, there's a link to 'Gene Association(s): TNPO3,IRF5, TNPO3, IRF5, HGVS Name(s): c.1864-1477T>C, c.1864-1477T>G, c.1870-1477T>C, c.1870-1477T>G, c.2062-1477T>C, c.2062-1477T>G, c.2164-1477T>C, c.2164-1477T>G'.

Ilustración 2: resultados de la búsqueda en Ensembl!

Nos redirige a una página (ilustración 3) de la que sacaremos diferente información, donde el apartado en **rojo** nos indica la naturaleza de esta variación (intrónica, codificante, ...) y pinchando sobre las consecuencias, nos redirige a éstas causadas por la mutación. El segundo apartado, en **naranja**, nos habla de la naturaleza de la variante, y de los alelos tanto ancestral como derivado, indicándonos la frecuencia del segundo alelo más común (MAF), y la más alta (en determinada población) de éste (*highest population* MAF). En este caso, concluimos que la variante es intrónica, donde el alelo ancestral es A (es el primero que aparece), derivado en C o G (que aunque aparezca último, no es el de menor frecuencia), donde ésta es de 0,06 para el alelo derivado (MAF) teniendo una frecuencia máxima de 0,22 (*highest population* MAF). Si dejamos el cursor sobre estas frecuencias, nos dice a que alelo se refiere.

rs12531711 SNP

Most severe consequence | [Intron variant](#) | [See all predicted consequences](#)

Alleles | [A/C/G](#) | Ancestral: [A](#) | MAF: [0.06](#) (G) | Highest population MAF: [0.22](#)

Location | [Chromosome 7:128617466](#) (forward strand) | VCF: [7 128617466 rs12531711 A C,G](#)

Evidence status ⓘ

HGVS names | This variant has [16](#) HGVS names - [Show](#) ⓘ

Synonyms | [Archive dbSNP](#) [rs17423567](#) ⓘ, [rs117836556](#) ⓘ, [rs60874173](#) ⓘ

Genotyping chips | This variant has assays on [12](#) chips - [Show](#) ⓘ

Original source | Variants (including SNPs and indels) imported from dbSNP (release 150) | [View in dbSNP](#) ⓘ

About this variant | This variant overlaps [5](#) transcripts, has [4102](#) sample genotypes, is associated with [5](#) phenotypes and is mentioned in [15](#) citations.

Description from SNPedia | association with [lupus](#) ⓘ [PMID:17412832 ⓘ] [PMID:21408207 ⓘ]... [Show](#) ⓘ

Explore this variant ⓘ

Genomic context

Genes and regulation

Flanking sequence

Population genetics

Phenotype data

Sample genotypes

Linkage disequilibrium

Phylogenetic context

Citations

Ilustración 3: identificación del SNP.

En el tercer apartado, en **amarillo**, observamos que la localización se encuentra en el cromosoma 7:128617466.

Echando un vistazo rápido por los demás apartados, en color **gris**, podemos consultar algunos datos más como sus SNPs sinónimos, o a cuantos transcritos afecta, además de los posibles genotipos o las citas en otras páginas.

Volvemos a incidir sobre la zona en **rojo**, sobre la cual haremos click para conocer los resultados:

Nos carga la siguiente información justo debajo del SNP que hemos localizado, sólo debemos movernos en la página hacia abajo (ilustraciones 4 y 5):

rs12531711 SNP

Most severe consequence | [Intron variant](#) | [See all predicted consequences](#)

Alleles | [A/C/G](#) | Ancestral: [A](#) | MAF: [0.06](#) (G) | Highest population MAF: [0.22](#)

Location | [Chromosome 7:128617466](#) (forward strand) | VCF: [7 128617466 rs12531711 A C,G](#)

Evidence status ⓘ

HGVS names | This variant has [16](#) HGVS names - [Show](#) ⓘ

Synonyms | [Archive dbSNP](#) [rs60874173](#) ⓘ, [rs17423567](#) ⓘ, [rs117836556](#) ⓘ

Genotyping chips | This variant has assays on [12](#) chips - [Show](#) ⓘ

Original source | Variants (including SNPs and indels) imported from dbSNP (release 150) | [View in dbSNP](#) ⓘ

About this variant | This variant overlaps [5](#) transcripts, has [4102](#) sample genotypes, is associated with [5](#) phenotypes and is mentioned in [15](#) citations.

Ilustración 4: detalle de la entrada a la descripción de transcritos.

Gene	Transcript (strand)	Allele (transcript allele)	Consequence Type
ENSG00000064419	ENST00000265388.5 (-)	C	Intron variant
HGNc: TNPO3	biotype: protein_coding	(G)	
ENSG00000064419	ENST00000265388.5 (-)	G	Intron variant
HGNc: TNPO3	biotype: protein_coding	(C)	
ENSG00000064419	ENST00000393245.1 (-)	C	Intron variant
HGNc: TNPO3	biotype: protein_coding	(G)	
ENSG00000064419	ENST00000393245.1 (-)	G	Intron variant
HGNc: TNPO3	biotype: protein_coding	(C)	
ENSG00000064419	ENST00000471166.1 (-)	C	Intron variant
HGNc: TNPO3	biotype: protein_coding	(G)	
ENSG00000064419	ENST00000471166.1 (-)	G	Intron variant
HGNc: TNPO3	biotype: protein_coding	(C)	
ENSG00000064419	ENST00000482320.1 (-)	C	Intron variant
HGNc: TNPO3	biotype: protein_coding	(G)	
ENSG00000064419	ENST00000482320.1 (-)	G	Intron variant
HGNc: TNPO3	biotype: protein_coding	(C)	

Ilustración 5: descripción de transcritos.

Aquí vemos que la mutación pertenece a la hebra molde, que afecta al gen TPNO3, y al transcrito ENSG00000064419. Nos muestra cada transcrito y su hebra complementaria

(podemos saberlo porque sus bases lo son), consultando el SNP de cada una, donde nos confirma que es una variante intrónica. Vemos también que nuestro transcrito codifica para una proteína. Nota: vemos que no hay correlaciones con expresiones del gen, ni EQTL (expresión cuantitativa de esta variable), ni relación con factores de regulación,... Esto puede ser porque como hemos dicho es intrónica.

Volvemos atrás en la página, y nos sale de nuevo la presentación de nuestro SNP (ilustración 6).

Comencemos ahora a indagar sobre las imágenes contenidas en el rectángulo verde. Nos centraremos sobre los apartados de '*Phenotype data*' y '*Population genetics*', rodeadas en azul. Simplemente, la información relacionada con los demás apartados que sea de nuestro interés, la consultaremos en otras páginas de forma diferente más adelante para obtener un estudio más completo. A su vez, nos servirá para contrastar información de esta página con otras, dándonos mayor criterio a la hora de seleccionarla.



Ilustración 6: identificación del SNP.

Si pulsamos sobre el dibujo de los fenotipos, nos da los siguientes datos (ilustración 7): nos muestra la patología a la que se asocia, el recurso de donde procede esta información, además de otros como el gen responsable y la probabilidad de tener este SNP concreto.

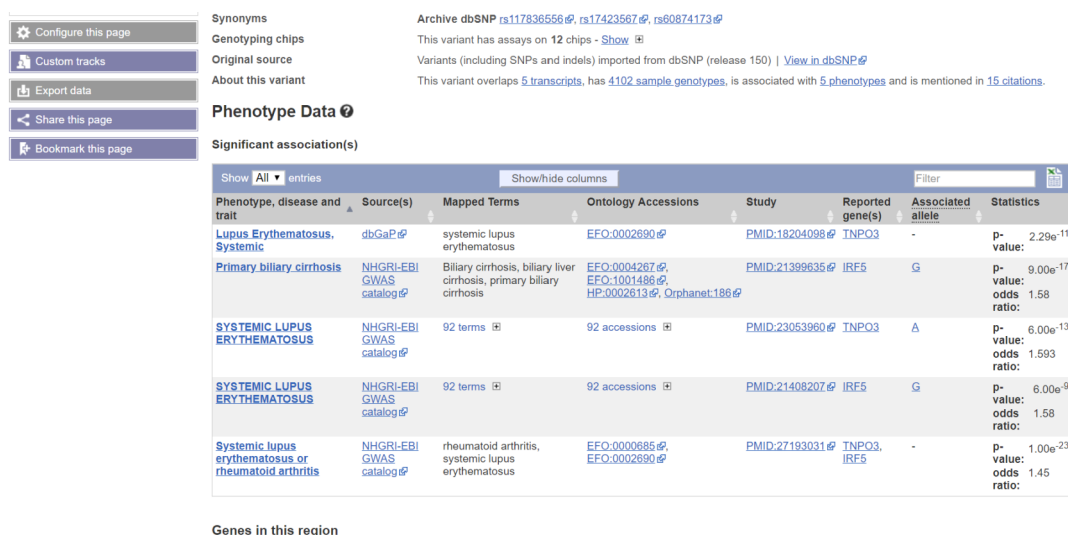


Ilustración 7: detalle de los fenotipos.

Vemos que se asocia con lupus eritematoso, cirrosis biliar, y artritis reumatoide. A la derecha de todo, en las estadísticas, vemos la probabilidad de tener este rasgo asociado con la variante derivada. También nos ofrece otros datos sobre los genes en los que influye, los recursos de los que han sido extraídos, o el alelo asociado, por ejemplo. La medida correspondiente a ‘odds ratio’ viene significando ‘razón de momios’ o de probabilidades, la cual nos relaciona en una población la posibilidad de manifestarse un fenotipo concreto de los que vemos en la tabla frente al tener o no un genotipo con el SNP en cuestión degenerado. Vemos que en todos los casos es mayor que uno, por lo que en términos estadísticos quiere decir que hay una correlación clara entre padecer una de las enfermedades mencionadas y contener la mutación estudiada.

Ahora volveremos atrás en la página para pinchar sobre el icono de genética de poblaciones.

Podemos de un vistazo rápido, hacernos una idea de cómo se reparten nuestros SNPs en el mundo, gracias a datos del proyecto 1000 Genomas (ilustración 8):

Population genetics

1000 Genomes Project Phase 3 allele frequencies

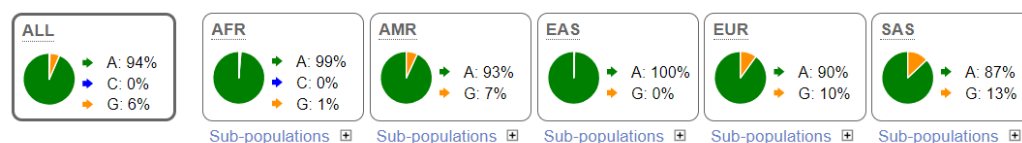


Ilustración 8: genética poblacional.

Nota: Para ver en detalle las poblaciones y subpoblaciones de todos los SNPs, consultar el Anexo I.

Ahora visitaremos la página de Svante Paäbo. Esto lo haremos para poder conocer los

alineamientos de nuestro Neanderthal y Vindija, la prueba arqueológica hallada en Europa de que este homínido efectivamente estuvo en el susodicho continente. Estos alineamientos nos permiten confirmar cuál era el alelo ancestral y cuál el derivado. Para ello, entraremos en la página: <https://www.eva.mpg.de/neandertal/index.html> Visualizaremos una página con el aspecto de la ilustración 9, y pincharemos sobre el buscador señalado:

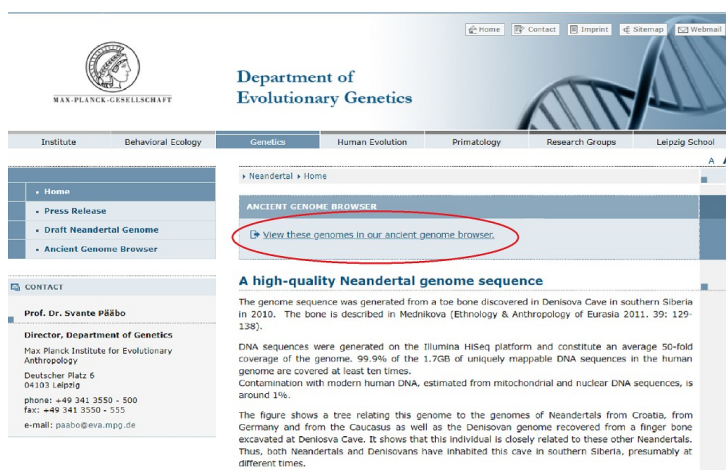


Ilustración 9: inicio del buscador de Svante Pääbo.

Entonces nos redirige a un buscador de alineamientos, donde debemos hacer dos cosas, la primera es poner las coordenadas de nuestro SNP (del ensamblaje GRCh37, que nos lo facilita la página del NCBI) buscando donde señala la flecha de la ilustración 10:

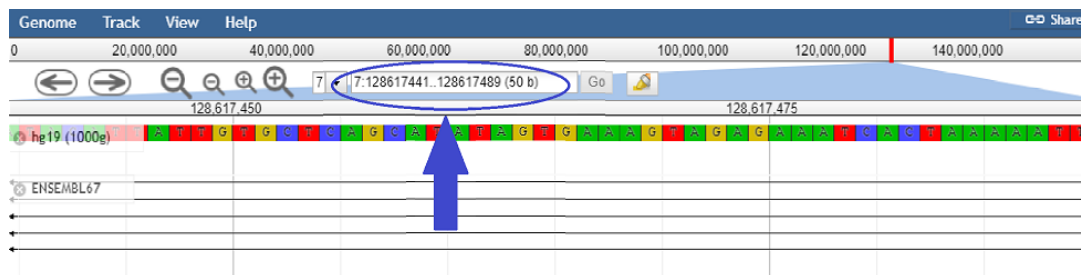


Ilustración 10: detalle de coordenadas introducidas en el buscador.

Y por otra parte, debemos asegurarnos de que estén los alineamientos que nos interesan, que son el Neanderthal y el Vindija, desmarcando las casillas que no nos interesen en la parte izquierda de la pantalla como muestra la ilustración 11.

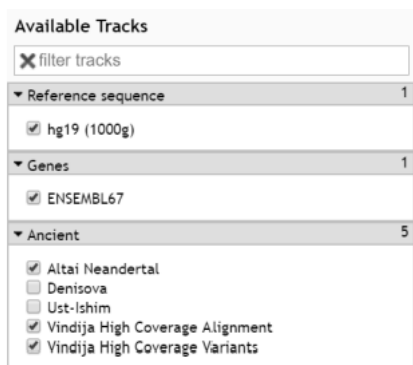


Ilustración 11: detalle de casillas marcadas y desmarcadas.

Obtendremos un alineamiento tal como se muestra en la ilustración 12:



Ilustración 12: alineamiento resultante.

Donde nos marca claramente nuestro SNP, señalando que el ancestral era A y ha sido derivado en G justo debajo de los rectángulos, marcando la posición del SNP. Vemos que todos los alineamientos se corresponden con G.

Ahora, procederemos a usar la página del NCBI:

<https://www.ncbi.nlm.nih.gov/SNP/>

Escribimos en nombre de nuestro SNP en el buscador (señalado con un círculo en la ilustración 13) y ejecutamos:

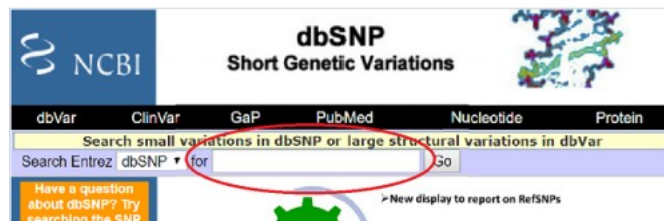


Ilustración 13: buscador del NCI.

Nos aparece una pantalla como la que se muestra en la ilustración 14, donde haremos click sobre el SNP de nuestro interés:

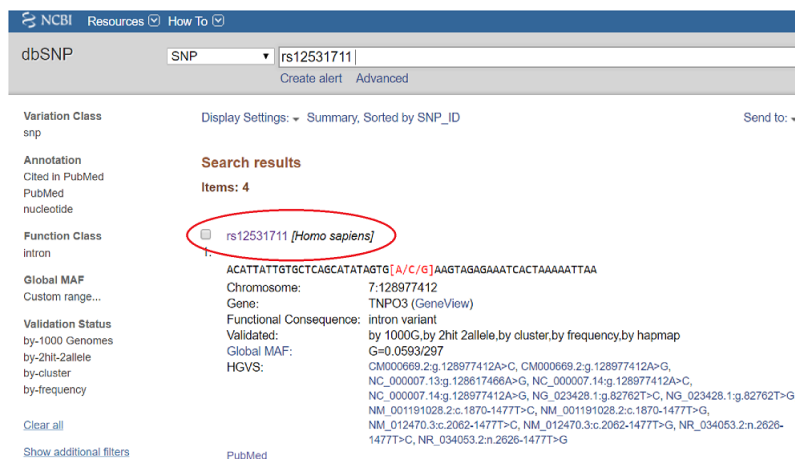


Ilustración 14: resultado de la búsqueda del SNP.

Nos redirigirá a una pantalla (ilustración 15) donde podremos ver situado nuestro SNP de la siguiente forma:



Ilustración 15: situación de nuestro SNP.

Está marcado en color verde, y nos indica el alelo ancestral primero (A) y luego los derivados (C/G). Podemos ver la secuencia flanqueante, así como ver el gen en el que nos estamos moviendo. Esta información es interesante porque, ya que hay secuencias que se heredan juntas, podemos relacionar nuestras mutaciones con otros fenotipos. Ahora entraremos en la página web de la Universidad de California, y usaremos su

buscador para continuar con el estudio de nuestro SNP.

Siguiendo el enlace: <https://genome.ucsc.edu/> se mostrará una pantalla como la ilustración 16:

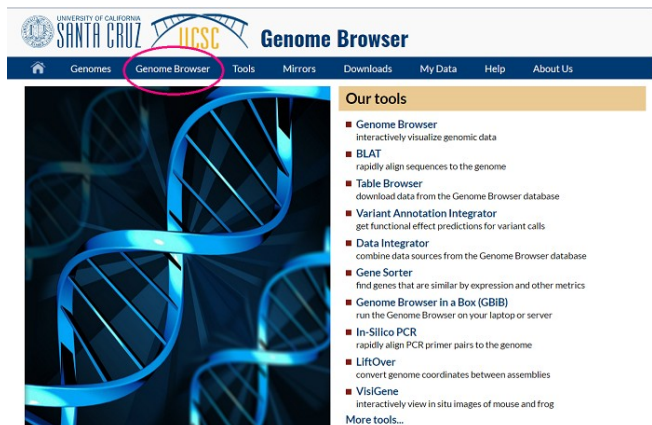


Ilustración 16: inicio del buscador de la UCSC.

Y vemos una pantalla (ilustración 17):

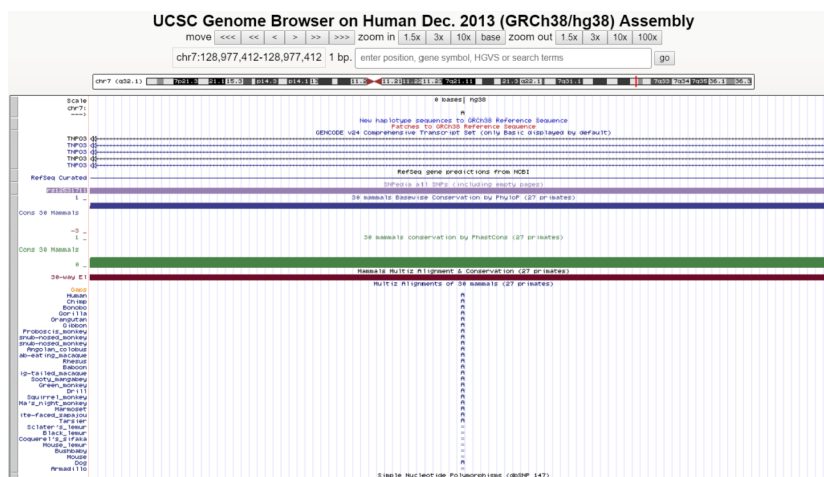


Ilustración 17: Buscador de huellas moleculares en ensamblaje hg38.

Podemos leer en la parte superior que por defecto, este buscador utiliza el ensamblaje GRCh38/hg38.

Si bajamos por la pantalla, nos encontramos todas estas opciones de configuración (ilustración 18):

+	Mapping and Sequencing	refresh
+	Genes and Gene Predictions	refresh
+	Phenotype and Literature	refresh
+	mRNA and EST	refresh
+	Expression	refresh
+	Regulation	refresh
+	Comparative Genomics	refresh
+	Variation	refresh
+	Repeats	refresh

Ilustración 18: opciones de configuración

Las desplegaremos y configuraremos de la siguiente manera, introduciendo además nuestro SNP en el buscador en la parte superior de la pantalla (ilustraciones 19 a la 23):

Mapping and Sequencing					
Base Position dense ▼	GRC Patch Release hide ▼	Alt Map... hide ▼	Assembly hide ▼	Centromeres hide ▼	Chromosome Band hide ▼
Clone Ends hide ▼	FISH Clones hide ▼	Gap hide ▼	GC Percent hide ▼	GRC Contigs hide ▼	GRC Incident hide ▼
Hg19 Diff hide ▼	INSDC hide ▼	LRG Regions hide ▼	Mappability... hide ▼	RefSeq Acc hide ▼	Restr Enzymes hide ▼
Scaffolds hide ▼	Short Match hide ▼	STS Markers hide ▼			

Ilustración 19: tutorial de búsqueda.

Genes and Gene Predictions					
GENCODE v24 pack ▼	NCBI RefSeq dense ▼	Other RefSeq hide ▼	All GENCODE... hide ▼	AUGUSTUS hide ▼	CCDS hide ▼
CRISPR... hide ▼	Geneid Genes hide ▼	Genscan Genes hide ▼	IKMC Genes Mapped hide ▼	LRG Transcripts hide ▼	MGC Genes hide ▼
Non-coding RNA... hide ▼	Old UCSC Genes hide ▼	ORFeome Clones hide ▼	Pfam in UCSC Gene hide ▼	RetroGenes V9 hide ▼	SGP Genes hide ▼
SIB Genes hide ▼	TransMap... hide ▼	UCSC Alt Events hide ▼	UniProt hide ▼		

Ilustración 20: tutorial de búsqueda.

Phenotype and Literature					
OMIM Alleles hide ▼	Cancer Gene Expr... hide ▼	ClinGen CNVs hide ▼	ClinVar Variants hide ▼	Coriell CNVs hide ▼	COSMIC Regions hide ▼
Development Delay hide ▼	Gene Interactions hide ▼	GeneReviews hide ▼	GWAS Catalog hide ▼	HGMD Variants hide ▼	OMIM Genes hide ▼
OMIM Pheno Loci hide ▼	SNPedia full ▼	UniProt Variants hide ▼			

mRNA and EST					
Human mRNAs hide ▼	Spliced ESTs hide ▼	Human ESTs hide ▼	Other ESTs hide ▼	Other mRNAs hide ▼	SIB Alt-Splicing hide ▼

Expression					
GTEx Gene hide ▼	GTEx Transcript hide ▼	Affy GNF1H hide ▼	Affy U133 hide ▼	Affy U95 hide ▼	GNF Atlas 2 hide ▼
GWIPS-viz Riboseq hide ▼					

Ilustración 21: tutorial de búsqueda.

mRNA and EST					
Human mRNAs hide ▼	Spliced ESTs hide ▼	Human ESTs hide ▼	Other ESTs hide ▼	Other mRNAs hide ▼	SIB Alt-Splicing hide ▼

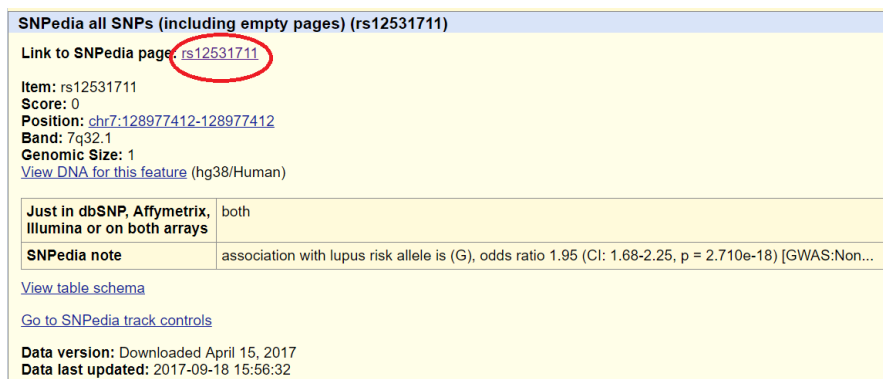
Expression					
GTEx Gene hide ▼	GTEx Transcript hide ▼	Affy GNF1H hide ▼	Affy U133 hide ▼	Affy U95 hide ▼	GNF Atlas 2 hide ▼
GWIPS-viz Riboseq hide ▼					

Regulation			
ENCODE Regulation... hide ▼	CpG Islands... hide ▼	ORRegAnno hide ▼	RefSeq Func Elems hide ▼

Comparative Genomics					
Conservation hide ▼	Cons 7 Verts hide ▼	Cons 20 Mammals hide ▼	Cons 30 Primates full ▼	Primate Chain/Net hide ▼	Placental Chain/Net hide ▼
Vertebrate Chain/Net hide ▼					

Ilustración 22: tutorial de búsqueda.

Nos redirige a otra pantalla, en la cual haremos de nuevo click sobre nuestro SNP (ilustración 26):



SNPedia all SNPs (including empty pages) (rs12531711)

Link to SNPedia page: [rs12531711](#)

Item: rs12531711
Score: 0
Position: chr7:128977412-128977412
Band: 7q32.1
Genomic Size: 1
[View DNA for this feature](#) (hg38/Human)

Just in dbSNP, Affymetrix, Illumina or on both arrays	both
SNPedia note	association with lupus risk allele is (G), odds ratio 1.95 (CI: 1.68-2.25, p = 2.710e-18) [GWAS:Non...

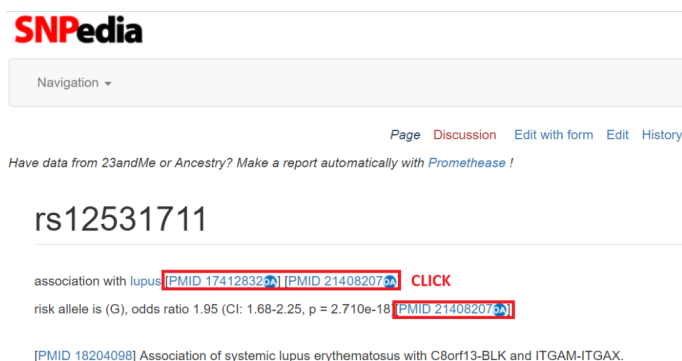
[View table schema](#)

[Go to SNPedia track controls](#)

Data version: Downloaded April 15, 2017
Data last updated: 2017-09-18 15:56:32

Ilustración 26: detalle enlace a artículos de la SNPedia.

Aquí vemos una nueva pantalla (ilustración 27), donde nos proporciona información sobre la asociación entre nuestra variante G y el lupus. Hacemos click sobre la referencia bibliográfica, donde leeré el *abstract* y sacaré mis propias conclusiones para cada SNP.



SNPedia

Navigation ▾

Page Discussion Edit with form Edit History

Have data from 23andMe or Ancestry? Make a report automatically with Promethease !

rs12531711

association with lupus [\[PMID 17412832\]](#) [\[PMID 21408207\]](#) **CLICK**

risk allele is (G), odds ratio 1.95 (CI: 1.68-2.25, p = 2.710e-18) [\[PMID 21408207\]](#)

[\[PMID 18204098\]](#) Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX.

Ilustración 27: resultado en SNPedia.

Concluimos que este SNP se encuentra en el locus IRF5, relacionado con la enfermedad de Lupus Eritematoso, siendo una enfermedad en la cual otros factores como los ambientales sugieren una asociación locus - enfermedad más débil, pero que no debe pasarse por alto (Chung et al., 2011).

Esta recopilación de información se hará exactamente igual para cada SNP.

Nota: también he hecho un seguimiento de la actividad de cada SNP en GTEx, pero ante la falta de espacio, para ver en detalle los alineamientos, identificación en el NCBI y su actividad, consultar Anexo II.

6. Resultados.

La genética evolutiva ha probado que los seres humanos modernos no africanos se han cruzado con el Neanderthal, y aunque apenas haya una huella genética de una centena de kilobases, es suficiente para detectarlo en locus específicos.

En un estudio realizado a más de mil personas en la actualidad, revela que la abundancia

de alelos del Neanderthal se encuentra en genes que afectan a los filamentos de queratina (componente principal del pelo y uñas), por lo que sugiere que han ayudado a adaptarse a ambientes no africanos.

Asimismo, se cree que la asociación de estos alelos con diferentes enfermedades podrían deberse a deriva genética, por lo que las zonas donde hay escaso material genético Neanderthal se encuentran enriquecidas en genes, y han causado esterilidad en mujeres, además de haberse reducido considerablemente en el cromosoma X (Patterson et al., 2014)

Por otra parte, se estima que conservamos cerca de un 2% del material genético procedente del Neanderthal, dándose una clara reducción de ese ADN en el cromosoma en los testículos frente a cualquier otro tejido.

Las mutaciones producidas por el material genético Neanderthal son altamente deletéreas ya que la población era mucho menor en el momento del cruce. Además, una esterilidad asociada con la hibridación entre Neanderthal-hombre arcaico puede ser la explicación a la importante reducción de mutaciones en el cromosoma X y en testículos, ya antes mencionado (Sankararaman, Mallick, Patterson, & Reich, 2016)

Los alelos utilizados en este estudio son 9, los cuales han sido identificados individualmente:

- SNP rs12531711. La variante es intrónica, donde el alelo ancestral es A (es el primero que aparece), derivado en C o G (que aunque aparezca último, no es el de menor frecuencia), donde ésta es de 0,06 para el alelo derivado (MAF) teniendo una frecuencia máxima de 0,22 (*Highest population MAF*). La localización se encuentra en el cromosoma 7:128617466. Concluimos que este SNP se encuentra en el loci IRF5, susceptible de enfermedad de Lupus Eritematoso, siendo una enfermedad en la cual otros factores como los ambientales sugieren una asociación locus - enfermedad más débil, pero que no debe pasarse por alto (Chung et al., 2011)
- SNP rs3025343. Corresponde a una región exónica no codificante. El SNP ancestral era G, actualmente puede ser sustituido por A, con una frecuencia lo más alta de 0,03, donde la población de mayor índice con este alelo es de 0,12. Se encuentra en el cromosoma 9:136478355, hebra molde. Hay evidencias consistentes aunque indirectamente relacionadas con el tabaquismo. Se han encontrado loci relacionados con el número de cigarrillos fumados al día, con el inicio de la práctica de fumar, y en el caso de nuestro SNP, se encuentra una asociación significativa con el receso de fumar (Tobacco & Consortium, 2010)
- SNP rs 7076156. El alelo ancestral era G, actualmente derivado en A, siendo esta

frecuencia de 0,13 en A (de media global), donde la frecuencia mayor se encuentra en un 0,32. Se trata de una mutación no - sinónima, en el cromosoma 10:62655424. De hecho la asociación más significativa entre el gen ZNF365 se da con ese SNP, que de hecho anula otras variaciones. Se estudiaron cuatro isoformas del gen ZNF365, y sólo en el cual se localiza este SNP (en una región exónica), vemos una reducción de la actividad de este gen en células transformadas por el virus de Epstein-Barr, la cual se asocia potencialmente con la enfermedad de Crohn (Vanoli et al., 2017), y se cree también que esta mutación afecta a la expresión de otros genes como ARL4A, MKKS, RRAGD, SUMF2, TDR1 y ZNF148 en enfermos (Haritunians et al., 2011)

· SNP rs12571093. Variante de la cadena positiva (*forward strand*), el alelo ancestral es el G, derivado en A, y siendo su frecuencia de 0,14. Su mayor frecuencia es de 0,49. Se encuentra en el cromosoma 10:70019371, y se asocia fenotípicamente con la medida del nervio óptico (área del disco). Muchos estudios asocian la enfermedad de Crohn con loci, incluidos aquellos con regiones no codificantes intergénicas de un sólo nucleótido polimórfico. No encontramos este SNP concreto en el texto pero sí lo localizamos en el gen ATOH7 en la página del NCBI (ilustración 28), y vemos que se encuentra en esta misma región:

Chromosome 10: 68,259,564-68,259,664

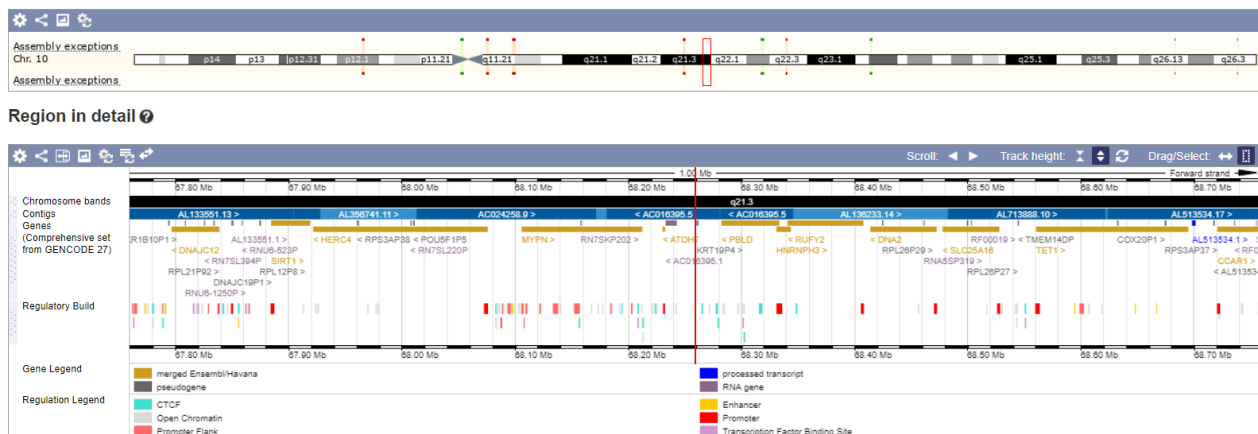


Ilustración 28: detalle de la localización de nuestro SNP en NCBI

Y este gen antes mencionado sí tiene una crucial relación con la media del área del disco óptico, además de tener un papel clave en las células que se encuentran en la retina. (Macgregor et al., 2010)

· SNP rs1834481. Variante exónica no codificante, el alelo ancestral era C substituído por G actualmente, siendo la frecuencia de 0,07 para G y con un 0,25 en la frecuencia de la población más alta. Se encuentra en el cromosoma 11:112023827 de la hebra molde. La interleucina-18 es un mediador de la inflamación relacionada con el desarrollo de

enfermedades cardiovasculares. Tras hacer una asociación en una población básicamente alemana, contrastando sujetos con infarto agudo de miocardio y el grupo control, se concluye que nuestro SNP forma un haplotipo con motivo TTAGC o bien GCGCA con efectos protectores frente a esta patología, (rs1946519-T, rs360717-T, rs5744241-A, rs1834481-G y rs3882891-C) sugieren que el gen de la interleucina 18 es un locus de susceptibilidad para el infarto agudo de miocardio, un hallazgo de posible interés en la práctica clínica. (Koch, Wolferstetter, Schatke, Schömig, & Kastrati, 2011)

- SNP rs11175593. Variante intrónica, alelo ancestral C substituído por T, donde la frecuencia de ésta es 0,05 y en la población donde la frecuencia es mayor, es 0,34. Se encuentra en el cromosoma 12:40601940, hebra molde.

- SNP rs75493593. Es una variante no - sinónima, la cual cambia el triplete resultando en un aminoácido diferente, pero sin modificar la longitud de la cadena. El alelo ancestral era G, siendo la segunda base más frecuente la T, que tiene una frecuencia de 0,06, y en la población de mayor frecuencia se encuentra en un 0,44. También puede ser substituída por C. Se encuentra en el cromosoma 17:6945087, hebra molde. En la enfermedad de Crohn (EC) se han identificado asociaciones con el polimorfismo de nucleótido único (SNP) rs11175593, pero no está claro cuál de los genes es el responsable: MUC19 o LRRK2. Tres SNP en MUC19 se asociaron significativamente con EC, pero no lo relaciona directamente con nuestro SNP. (Kumar et al., 2013)

- SNP rs75418188. Nuevamente mutación sin sentido, donde la base con la segunda frecuencia más alta es la T, con un 0,06, y la población con la frecuencia más alta cuenta con un 0,44. El alelo ancestral es C, y también pudo derivarse en G. Se encuentra en el cromosoma 17:6945483.

- SNP rs117767867. También variante no - sinónima, alelo ancestral C derivado en T con una frecuencia de 0,06 siendo la frecuencia más alta de 0,44. Se encuentra nuevamente en el cromosoma 17:6946330.

Para estos tres últimos alelos, no he encontrado evidencias claras de asociación a patologías.

Poblacionalmente, todos los alelos siguen el mismo tipo de distribución, donde en África no se encuentran nunca los alelos derivados, prueba de que el cruce con el Neanderthal se produjo una vez el hombre moderno salió de susodicho continente.

En cuanto a la localización cromosómica, todos los SNPs están efectivamente fuera del cromosoma X y ninguno relacionado con genes de expresión en el tejido de los testículos. Y las huellas moleculares han mostrado en este estudio que los alelos derivados en todo caso son propios del Neanderthal. Es decir, la huella molecular muestra que todos los

alelos del Neanderthal son mutaciones posteriores a la separación del linaje humano.

7. Discusión.

Todos estos resultados sugieren una pregunta, ¿por qué no han sido eliminados estos alelos si a veces son perjudiciales para el ser humano?

Bueno, sabemos que la población Neanderthal era mucho más pequeña que la de los humanos, así que los efectos de la deriva genética eran más intensos. Esto significa que los alelos de los Neanderthales se han fijado en su genoma no por selección natural, sino por cambios aleatorios (Juric, Aeschbacher, & Coop, 2016)

Si la mayor parte de las mutaciones deletéreas que han aparecido recientemente en poblaciones humanas fuesen generalmente recesivas, la fracción Neanderthal en nuestro genoma (generalmente también deletérea) podría ofrecer un efecto protector, dado que las posiciones afectadas por esas mutaciones deletéreas en alelos de ambas especies son rara vez las mismas (Harris & Nielsen, 2016)

Además, la presencia de haplotipos Neanderthales puede ser beneficiosa para los humanos porque algunos de estos genes son protectores ante otras enfermedades, o estimulan el sistema inmunológico para defenderse contra los hongos, parásitos, así como bacterias. Entonces, cuando los humanos colonizaron nuevos ambientes, cogieron un 'atajo genético', porque estos genes beneficiosos ya estaban allí. Y sabemos que éstos fueron útiles en el pasado, pero ahora, los ambientes y el estilo de vida son diferentes. Por ejemplo, eran beneficiosos para personas que viven en malas condiciones, pero también son responsables de muchos casos de alergias, respuesta inflamatoria o autoinmune.

Otro ejemplo, es un gen que hace que la sangre tenga más facilidad para coagular. Cuando nuestros antecesores cazaban o daban a luz a niños con un cráneo grande, para no desangrarse y sobrevivir esto era muy importante, pero ahora, vivimos por muchos más años y este factor es un riesgo que incrementa los derrames y coágulos cerebrales (Gibbons, 2016)

Discussion:

All these results suggest a question, why have these alleles not been eliminated if sometimes they are harmful to the human being?

Good, we know that the population Neanderthal was much smaller than that of the human beings, so the effects of the genetic drift were more intense. This means that the alleles of the Neanderthales have been fixed in their genome not by natural selection, but by

random changes (Juric et al., 2016)

If most of the deleterious mutations that have appeared recently in human populations were generally recessive, the fraction Neanderthal in our genome (generally also deleterious) might offer a protective effect, since the positions affected by those deleterious mutations in alleles of both species are rarely them (Harris & Nielsen, 2016). Besides, the presence of haplotypes Neanderthals can be beneficial to the human beings because some of these genes are protectors in the face of other illnesses, or stimulate the immune system to be defended against the fungi, parasitic, as well as bacteria. Then, when the human beings colonized new environments, they caught a 'genetic shortcut', because these beneficial genes were already there. And we know that these were useful in the past, but now, the environments and the way of life are different. For example, they were beneficial to people who live in bad conditions, but are also responsible for many cases of allergies, inflammatory or auto-immune answer.

Another example, is a gene that makes the blood have more easiness to coagulate. When our predecessors hunted or gave birth to children with a big skull, in order not to bleed and to survive this was very important, but now, we live for many more years and this factor is a risk that increases the spillings and cerebral coagula (Gibbons, 2016)

8.Conclusiones.

1. La cantidad mínima estimada de introgresión procedente del Neanderthal es de un 2%, aunque se sospecha que haya muchos alelos más con este origen.
 2. Actualmente existe una situación controvertida debido a que no se acaba de comprender por qué estos alelos no se han eliminado totalmente de nuestro genoma, ya que muchos de ellos parecen ser parcialmente deletéreos.
 3. Se ha confirmado el origen introgresivo de 9 alelos en otros tantos Polimorfismos de Nucleótido Único (SNPs) candidatos.
 4. De estos 9 SNPs, he encontrado evidencias de patologías en estos casos:
rs12531711: Lupus eritematoso
rs3025343: Tabaquismo
rs 7076156: Enfermedad de Crohn.
rs12571093: Área del disco óptico (interfiere con el gen ATOH7).
rs1834481: Enfermedades cardiovasculares e infarto agudo de miocardio en especial (debido a que modifica los niveles de interleucina-18).
- Para el resto de alelos, no he encontrado evidencias claras de asociación a otras patologías.

5. En cuanto a las huellas moleculares de la selección, hay, por un lado, una clara asimetría en la distribución cromosómica de las regiones introgresadas, con una ausencia completa de introgresiones ligadas al cromosoma X. Por otro lado, en los 9 SNPs examinados en este estudio el alelo ancestral es siempre el de la especie humana, mientras que el derivado es el procedente del Neanderthal, siendo éste último siempre el minoritario.
6. Estas huellas moleculares son prueba de una intensa acción de la selección natural contra los alelos procedentes del Neanderthal.
7. La elevada frecuencia en el genoma Neanderthal de alelos que son hoy en día aparentemente deletéreos podría deberse a la deriva genética, ya que la población Neanderthal se encontraba en un estado de cuello de botella donde el azar tuvo más impacto que la selección natural. Estos alelos deletéreos pudieron acelerar su extinción.
8. Es posible también que estos alelos fueran beneficiosos para el Neanderthal para las particulares condiciones ambientales en las que se encontraban sus poblaciones.
9. Cuando el hombre moderno durante su expansión por Eurasia se encontró con el Neanderthal, incorporó de esta manera un 'atajo genético' a la hora de colonizar nuevos ambientes. Esas condiciones ambientales del pasado no persisten en la actualidad, y es posible que no se hayan eliminado estos alelos debido a que no haya transcurrido el tiempo suficiente.

Conclusions:

1. The estimated minimum amount of introgression from the Neanderthal is 2%, although it is suspected that there are many more alleles with this origin.
2. There is currently a controversial situation because we do not understand why these alleles have not been completely eliminated from our genome, since many of them seem to be partially deleterious.
3. The introgressive origin of 9 alleles has been confirmed in as many other candidate Single Nucleotide Polymorphisms (SNPs).
4. Of these 9 SNPs, I have found evidence of pathologies in these cases:
rs12531711: Lupus erythematosus
rs3025343: Smoking behavior.
rs 7076156: Crohn's disease.
rs12571093: Optic disc area (interferes with the ATOH7 gene).
rs1834481: Cardiovascular diseases and myocardial infarction (because this SNP modifies the levels of interleukin-18).

For the other alleles, I have not found clear evidence of association with other pathologies.

5. Regarding the molecular traces of the selection, there is, on the one hand, a clear asymmetry in the chromosomal distribution of the introgressed regions, with a complete absence of introgressions linked to the X chromosome. On the other hand, in the 9 SNPs examined in this study, the ancestral allele is always from the ancestral human specie, while the derivative is that derived from the Neanderthal, the latter being always the minority.

6. These molecular traces are evidence of an intense action of natural selection against the alleles coming from the Neanderthal.

7. The high frequency in the Neanderthal genome of alleles that are now apparently deleterious could be due to genetic drift, since the Neanderthal population was in a bottleneck state where chance had more impact than natural selection. These deleterious alleles could accelerate their extinction.

8. It is also possible that these alleles were beneficial for the Neanderthal for the particular environmental conditions in which their populations were found.

9. When modern man during his expansion through Eurasia met the Neanderthal, he incorporated a 'genetic shortcut' when colonizing new environments. Those environmental conditions of the past do not persist at present, and it is possible that these alleles have not been eliminated because sufficient time has not elapsed.

9. Bibliografía.

Ackermann, R. R., Mackay, A., & Arnold, M. L. (2016). The Hybrid Origin of "Modern" Humans. *Evolutionary Biology*. <https://doi.org/10.1007/s11692-015-9348-1>

Chung, S. A., Taylor, K. E., Graham, R. R., Nititham, J., Lee, A. T., Ortmann, W. A., ... Criswell, L. A. (2011). Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. *PLoS Genetics*. <https://doi.org/10.1371/journal.pgen.1001323>

Gibbons, A. (2016). Neandertal genes linked to modern diseases. *Science*. <https://doi.org/10.1126/science.351.6274.648>

Haritunians, T., Jones, M. R., McGovern, D. P. B., Shih, D. Q., Barrett, R. J., Derkowski, C., ... Taylor, K. D. (2011). Variants in ZNF365 isoform D are associated with Crohn's disease. *Gut*. <https://doi.org/10.1136/gut.2010.227256>

Harris, K., & Nielsen, R. (2016). The genetic cost of neanderthal introgression. *Genetics*. <https://doi.org/10.1534/genetics.116.186890>

Hogenboom, M. (2015). Why are we the only human species still alive? Retrieved from

<http://www.bbc.com/earth/story/20150929-why-are-we-the-only-human-species-still-alive>

- Juric, I., Aeschbacher, S., & Coop, G. (2016). The Strength of Selection against Neanderthal Introgression. *PLoS Genetics*.
<https://doi.org/10.1371/journal.pgen.1006340>
- Kelso, J., & Prüfer, K. (2014). Ancient humans and the origin of modern humans. *Current Opinion in Genetics and Development*. <https://doi.org/10.1016/j.gde.2014.09.004>
- Koch, W., Wolferstetter, H., Schatke, A., Schömig, A., & Kastrati, A. (2011). Interleukin 18 gene variation and risk of acute myocardial infarction. *Cytokine*.
<https://doi.org/10.1016/j.cyto.2011.09.006>
- Kumar, V., Mack, D. R., Marcil, V., Israel, D., Krupoves, A., Costea, I., ... Levy, E. (2013). Genome-wide association study signal at the 12q12 Locus for Crohn's disease may represent associations with the MUC19 gene. *Inflammatory Bowel Diseases*.
<https://doi.org/10.1097/MIB.0b013e318281f454>
- Macgregor, S., Hewitt, A. W., Hysi, P. G., Ruddle, J. B., Medland, S. E., Henders, A. K., ... Mackey, D. A. (2010). Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. *Human Molecular Genetics*.
<https://doi.org/10.1093/hmg/ddq144>
- Patterson, N., Reich, D., Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., ... Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. <https://doi.org/10.1038/nature12961>
- Sankararaman, S., Mallick, S., Patterson, N., & Reich, D. (2016). The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Current Biology*. <https://doi.org/10.1016/j.cub.2016.03.037>
- Sankararaman, S., Patterson, N., Li, H., Pääbo, S., & Reich, D. (2012). The Date of Interbreeding between Neandertals and Modern Humans. *PLoS Genetics*.
<https://doi.org/10.1371/journal.pgen.1002947>
- Tobacco, T., & Consortium, G. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature Genetics*. <https://doi.org/10.1038/ng.571>
- Vanoli, A., Di Sabatino, A., Martino, M., Dallera, E., Furlan, D., Mescoli, C., ... Solcia, E. (2017). Epstein-Barr virus-positive ileal carcinomas associated with Crohn's disease. *Virchows Archiv*. <https://doi.org/10.1007/s00428-017-2209-9>
- Varki, A. (2016). Why are there no persisting hybrids of humans with Denisovans, Neanderthals, or anyone else? Retrieved from
<http://www.pnas.org/content/113/17/E2354.full>